(5/10/2007)

TEMA 1: ERRORES.

1. INTRODUCCIÓN: RESOLUCIÓN APROXIMADA DE PROBLEMAS.

¿Por qué hacer una resolución aproximada y no exacta?

* Falta de métodos para resolver los problemas de forma exacta.
* Conocemos el método exacto, pero es una labor tediosa, pues se trata de problemas de grandes dimensiones.

  Ej: Un sistema de:

  $\left.\begin{array}{l}\text{1000 ecs.}\\ \text{1000 incógnitas}\end{array}\right\}$

* Es suficiente una aproximación (aunque el problema no sea excesivamente grande).

Para hacer este tipo de resoluciones utilizamos MÁQUINAS (ordenadores). Por ello, hay que tener en cuenta que las máquinas tienen recursos finitos, es decir, capacidad limitada.

EJ: Tenemos el número $x = 1'333\ldots$ Este número periódico lo podemos almacenar en forma de fracción: $\frac{4}{3}$

$$\begin{array}{rl} - \quad x = & 1'333\ldots \\ 10x = & 13'333\ldots \\ \hline 9x = & 12 \implies \frac{12}{9} = \frac{4}{3} = x \end{array}$$

$x = 1'2345\ldots101112\ldots$ → Este número no se puede expresar como fracción, así que no se puede almacenar.

Para almacenar números no racionales tendremos que almacenarlo de forma aproximada ($\tilde{x}$). Al hacerlo cometeremos un ERROR DE REDONDEO.

EJ: $x = 1'23456 \dots 1011 \dots$

4 decimales ... Tenemos 2 formas de hacerlo:

$$\left.\begin{array}{ll} \tilde{x} = 1'2345 & \text{(TRUNCAR)} \\ \approx\!\!\!x = 1'2346 & \text{(REDONDEAR)} \end{array}\right\} \Rightarrow \text{ERROR DE REDONDEO.}$$

$\Downarrow$

TRUNCAR $\neq$ REDONDEAR

pero en los 2 casos se llama error de redondeo (no se dice error de truncamiento)

Los errores de redondeo se van acumulando operación tras operación, por lo que al final podemos llegar a un resultado completamente diferente del esperado.

EJ: $127 = 1.100 + 2.10 + 7$

* REPRESENTACIÓN EN BASE DECIMAL DE ENTEROS POSITIVOS:

$$x = a_0 . 1 + a_1 . 10 + a_2 . 10^2 + \dots + a_n . 10^n \qquad a_i \in \{0, 1, \dots, 9\}$$

* REPRESENTACIÓN EN BINARIO:

$$x = b_0 . 2^0 + b_1 . 2^1 + b_2 . 2^2 + \dots + b_n . 2^n \qquad b_i \in \{0, 1\}$$

Se suele representar como:

$$b_n b_{n-1} b_{n-2} \dots b_1 b_0 \quad \text{ó} \quad \underbrace{b_0 b_1 \dots b_{n-1} b_n}_{\text{MÁS PRÁCTICO.}}$$

EJ: $100101 \longrightarrow 1.2^0 + 0.2^1 + 0.2^2 + 1.2^3 + 0.2^4 + 1.2^5$

Cualquier número se puede representar como:

$$x = \pm r \cdot 2^{e} \quad ; \quad \frac{1}{2} \leq r < 1 \qquad \text{BINARIO}$$

ó

$$x = \pm q \cdot 10^{p} \quad ; \quad \frac{1}{10} \leq q < 1 \qquad \text{DECIMAL}$$

donde:

e = exponente

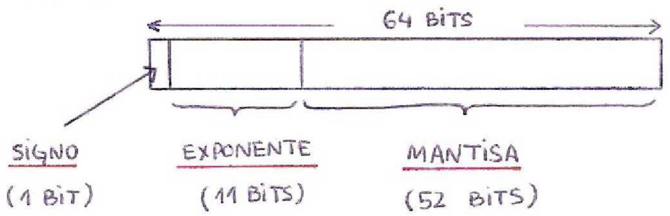
r = mantisa

$r \cdot 2^{e} = 2r \cdot 2^{e-1}$ ... pero se suele representar siempre como $r \cdot 2^{e}$

2. REPRESENTACIÓN DEL ESTÁNDAR IEEE.

$$\left.\begin{array}{l} e = \text{exponente} \\ r = \text{mantisa} \end{array}\right\} \; x = \pm r \cdot 2^{e}$$

Utilizamos 64 bits:



• SIGNO:

$0 \longrightarrow + \quad (-1)^{0}$

$1 \longrightarrow - \quad (-1)^{1}$

• EXPONENTE:

$$2^{11} \text{ posibilidades} \left\{ \begin{array}{l} 00 \ldots 000 \\ \\ 11 \ldots 111 \end{array} \right.$$

$2^{11} = 2^{10} \cdot 2 = 1024 \cdot 2 = 2048$ posibles exponentes.

Podemos disponer de 2048 exponentes... ¿cuáles nos convienen más? El estándar IEEE los reparte de forma simétrica alrededor del 0.

Rango del exponente $= \{-1023, \dots, 1024\} \ni e$

Exponente (11 bits) $\longrightarrow a_0 a_1 \dots a_{10} \quad ; \quad a_i \in \{0,1\}$

$$0 \leqslant \underbrace{a_0 2^0 + a_1 2^1 + a_2 2^2 + \dots + a_{10} 2^{10}}_{C} \leqslant \frac{1-2^{11}}{1-2} = 2^{11} - 1 = 2047$$

$$-1023 \leqslant \underbrace{C - 1023}_{e} \leqslant 1024$$

Los 11 bits a 1 me dan 2047, pero como quiero tener el rango $\{-1023, 1024\}$, le resto 1023 a los bits para obtener el exponente.

• MANTISA:

Son 52 bits : $1, b_1 b_2 b_3 \dots b_{52} \quad ; \quad b_i \in \{0,1\}$

¡NO EMPIEZA POR 0!!

Este 1 no se escribe, ya que la mantisa es 1'...

Los bits no empiezan desde el 0 porque quiero que la mantisa me represente:

$$m = \underbrace{b_1 \frac{1}{2} + b_2 \left(\frac{1}{2}\right)^2 + \dots + b_{52} \left(\frac{1}{2}\right)^{52}}$$

$$0 \leqslant m \leqslant \frac{\frac{1}{2} - \left(\frac{1}{2}\right)^{53}}{1 - \frac{1}{2}} = 1 - \left(\frac{1}{2}\right)^{52} < 1$$

$0 \leq m < 1$

$1 \leq m+1 < 2 \longrightarrow$

$m+1 = r$ (mi verdadera mantisa, la que quiero expresar). Esto viene de haber eliminado el 1 de la representación de la mantisa.

$\boxed{x = \pm r \cdot 2^e}$ REPRESENTACIÓN EN COMA FLOTANTE.

Si queremos representar un número con más de 52 bits en la mantisa, habrá que aproximar:

$x = \pm r \cdot 2^e$ $\qquad r = 1 + m$

sólo represento esto.

$r = b_0 b_1 b_2 \dots b_{83} \longrightarrow$ Sobran muchos bits. Tengo que aproximarlo porque sólo puedo representar 52.

Al aproximarlo, cometemos un ERROR DE REDONDEO, que puede ser de 2 tipos:

ERROR ABSOLUTO : $|x - \tilde{x}|$

ERROR RELATIVO : $\dfrac{|x-\tilde{x}|}{|x|}$

donde $\begin{cases} x \Rightarrow \text{N}^{\text{o}} \text{ a representar} \\ \tilde{x} \Rightarrow \text{Aproximación} \end{cases}$

Métodos para aproximar x a $\tilde{x}$ $\begin{cases} \text{TRUNCAR} \\ \text{REDONDEAR} \end{cases}$

Para poder redondear el exponente, sabiendo que 2^{c-1023} donde $c - 1023 = e$ (exponente) tenemos:

$$\frac{1}{2^{1023}} \cdot 2^{-1023} \leq e \leq 2^{1024}$$

Si intentamos representar un nº mayor que 2^{1024} obtenemos OVERFLOW. Si intentamos representar un nº menor que 2^{-1023} obtenemos UNDERFLOW.

3. ÈPSILÓN DE LA MÁQUINA.

Nos dan un número x, pero este x no lo podemos representar. Por lo tanto, representamos $\tilde{x}$, donde:

error relativo

$$\tilde{x} = x.(1+\delta)$$
$$= x + x\delta$$

$$\boxed{|\delta| \leq \mathcal{E}}$$

Èpsilón de la máquina

$$|\tilde{x} - x| = |x\delta| \quad ; \quad \frac{|\tilde{x} - x|}{|x|} = |\delta|$$

δ => ERROR DE REPRESENTACIÓN

Ejercicio:

$$x = \pm(1+m)\,2^{e}$$

más de 52 bits

$$\tilde{x} = \pm(1+\tilde{m})\,2^{\tilde{e}}$$

sólo 52 bits

donde $e = \tilde{e}$ (Nos fijamos sólo en la mantisa).

$|m - \tilde{m}|$ ⟶ Necesitamos dar una cota para m.

(9/10/2007)

EJERCICIOS : TEMA 1 : ERRORES

* PROBLEMA 1.

Supongamos que los valores $\tilde{x}_1$ y $\tilde{x}_2$ aproximan a x_1 y x_2 $(x_1 \neq x_2)$ respectivamente, con un error relativo de $\mathcal{E}_1$ y $\mathcal{E}_2$ siendo $0 \leq \mathcal{E}_1, \mathcal{E}_2 \leq \mathcal{E}$

a) Determinar una cota de $\mathcal{E}$ que permita garantizar que $\tilde{x}_1 \neq \tilde{x}_2$.

b) La magnitud $\dfrac{1}{x_1 - x_2}$ se aproxima por $\dfrac{1}{\tilde{x}_1 - \tilde{x}_2}$ en las condiciones anteriores. Estimar el error relativo cometido en la aproximación, suponiendo $\mathcal{E}$ suficientemente pequeño.

EL ERROR ABSOLUTO DA POCA INFORMACIÓN, ES MEJOR EL RELATIVO.
(No es lo mismo cometer un error de 3 km. en una medida de 20 km. que de 1000 km.)

ⓐ $\begin{array}{ccc} \tilde{x}_1 & \sim & x_1 \\ & & \neq \\ \tilde{x}_2 & \sim & x_2 \end{array}$ ($\tilde{x}_1, \tilde{x}_2$ representables en una máquina porque son las aproximaciones).

$$\left| \frac{\tilde{x}_1 - x_1}{x_1} \right| = \mathcal{E}_1 \leq \mathcal{E}$$

El error relativo está acotado

$$\frac{|\tilde{x}_1 - x_1|}{|x_1|} = \mathcal{E}_1 \quad ; \quad |\tilde{x}_1 - x_1| = \mathcal{E}_1 . |x_1| \leq \mathcal{E} . |x_1| \iff$$

$$\iff -\mathcal{E} . |x_1| \leq \tilde{x}_1 - x_1 \leq \mathcal{E} |x_1| \overset{\text{suma } x_1}{\iff} x_1 - \mathcal{E} . |x_1| \leq \tilde{x}_1 \leq x_1 + \mathcal{E} . |x_1|$$

Esto mismo se puede hacer para $\mathcal{E}_2$ y obtenemos: $x_2 - \mathcal{E} . |x_2| \leq \tilde{x}_2 \leq x_2 + \mathcal{E} . |x_2|$

$$\left.\begin{array}{l} x_1 - \mathcal{E}|x_1| \leq \tilde{x}_1 \leq x_1 + \mathcal{E}|x_1| \\ x_2 - \mathcal{E}|x_2| \leq \tilde{x}_2 \leq x_2 + \mathcal{E}|x_2| \end{array}\right\} \text{ Ahora quiero acotar } \tilde{x}_1 - \tilde{x}_2$$

$$x_1 - \mathcal{E}|x_1| - \tilde{x}_2 \leq \tilde{x}_1 - \tilde{x}_2 \leq x_1 + \mathcal{E}|x_1| - \tilde{x}_2$$

ya que es $>$ que $\tilde{x}_2$ — ya que es menor que $\tilde{x}_2$

$$x_1 - \mathcal{E}|x_1| - \underbrace{(x_2 + \mathcal{E}|x_2|)} \qquad x_1 + \mathcal{E}|x_1| - \underbrace{(x_2 - \mathcal{E}|x_2|)}$$

$$\underline{x_1 - \mathcal{E}|x_1| - (x_2 + \mathcal{E}|x_2|)} \leq x_1 - \mathcal{E}|x_1| - \tilde{x}_2 \leq \underline{\tilde{x}_1 - \tilde{x}_2} \leq x_1 + \mathcal{E}|x_1| - \tilde{x}_2 \leq \underline{x_1 + \mathcal{E}|x_1| - (x_2 - \mathcal{E}|x_2|)}$$

Finalmente obtenemos:

$$\underline{(x_1 - x_2) - \mathcal{E}(|x_1| + |x_2|) \leq \tilde{x}_1 - \tilde{x}_2 \leq (x_1 - x_2) + \mathcal{E}(|x_1| + |x_2|)}$$

Buscamos que $\tilde{x}_1 \neq \tilde{x}_2$. Vamos a diferenciar 2 casos:

1º) Hacemos que $(x_1 - x_2) - \mathcal{E}(|x_1| + |x_2|) > 0$

Esto es: $\underbrace{\mathcal{E}(|x_1| + |x_2|)}_{\text{positivo}} < x_1 - x_2$

$$\boxed{\mathcal{E} < \frac{x_1 - x_2}{|x_1| + |x_2|} \quad ; \quad (x_1 > x_2)}$$ SOLUCIÓN 1^er^ CASO.

2º) Hacemos que $\underbrace{(x_1 - x_2)}_{\text{negativo}} - \underbrace{\mathcal{E}(|x_1| + |x_2|)}_{\text{positivo}} < 0$

Esto es: $\mathcal{E}(|x_1| + |x_2|) < x_2 - x_1$

$$\boxed{\mathcal{E} < \frac{x_2 - x_1}{|x_1| + |x_2|} \quad ; \quad (x_2 > x_1)}$$ SOLUCIÓN 2º CASO.

Si se cumple 1º) ó 2º) $\mathcal{E}$ no es 0.

Por tanto, para que $\tilde{x}_1 \neq \tilde{x}_2$ ha de cumplirse: $\boxed{\mathcal{E} < \dfrac{|x_2 - x_1|}{|x_1| + |x_2|}}$

ⓑ Cantidad a aproximar $\longrightarrow \dfrac{1}{x_1 - x_2}$

Cantidad aproximada $\longrightarrow \dfrac{1}{\tilde{x}_1 - \tilde{x}_2}$

$$\text{ERROR ABSOLUTO} = \left| \frac{1}{x_1 - x_2} - \frac{1}{\tilde{x}_1 - \tilde{x}_2} \right|$$

$$\text{ERROR RELATIVO} = \frac{\left| \frac{1}{x_1 - x_2} - \frac{1}{\tilde{x}_1 - \tilde{x}_2} \right|}{\left| \frac{1}{x_1 - x_2} \right|}$$

Sabemos que $x_1 \neq x_2$ y que $\tilde{x}_1 \neq \tilde{x}_2$ (por el apartado anterior).

$$\frac{\left| \frac{1}{x_1 - x_2} - \frac{1}{\tilde{x}_1 - \tilde{x}_2} \right|}{\left| \frac{1}{x_1 - x_2} \right|} = \left| \frac{\frac{1}{x_1 - x_2} - \frac{1}{\tilde{x}_1 - \tilde{x}_2}}{\frac{1}{x_1 - x_2}} \right| = \left| 1 - \frac{x_1 - x_2}{\tilde{x}_1 - \tilde{x}_2} \right| = (*)$$

Tenemos que: $|\tilde{x}_1 - x_1| = \mathcal{E}_1 |x_1|$ (apartado a).

$$\underset{(\mathcal{E}_1)}{x_1 - \mathcal{E}|x_1|} \leq \tilde{x}_1 \leq \underset{(\mathcal{E}_1)}{x_1 + \mathcal{E}|x_1|}$$

$$\underset{(\mathcal{E}_2)}{x_2 - \mathcal{E}|x_2|} \leq \tilde{x}_2 \leq \underset{(\mathcal{E}_2)}{x_2 + \mathcal{E}|x_2|}$$

$$(*) = \frac{|\tilde{x}_1 - \tilde{x}_2 - x_1 + x_2|}{|\tilde{x}_1 - \tilde{x}_2|} = \frac{|(\tilde{x}_1 - x_1) + (x_2 - \tilde{x}_2)|}{|\tilde{x}_1 + \tilde{x}_2|} \underset{\substack{\text{desigualdad} \\ \text{triangular}}}{\leq} \frac{|\tilde{x}_1 - x_1| + |\tilde{x}_2 - x_2|}{|\tilde{x}_1 - \tilde{x}_2|} =$$

$$= \frac{\mathcal{E}|x_1| + \mathcal{E}_2|x_2|}{|\tilde{x}_1 - \tilde{x}_2|} \leq \frac{\mathcal{E}(|x_1| + |x_2|)}{\underbrace{\underbrace{(x_1 - x_2)}_{>0} - \mathcal{E}(|x_1| + |x_2|)}_{>0}} \quad ; \quad x_1 > x_2 \quad \tilde{x}_1 > \tilde{x}_2$$

(Tenemos del apartado a.) $(x_1 - x_2) - \mathcal{E}(|x_1| + |x_2|) \leq \tilde{x}_1 - \tilde{x}_2 \leq (x_1 - x_2) - \mathcal{E}(|x_1| + |x_2|)$

Podríamos dejarlo como alguna de estas 2 formas equivalentes:

$$\boxed{\frac{\varepsilon(|x_1|+|x_2|)}{(x_1-x_2)-\varepsilon(|x_1|+|x_2|)}} \qquad \boxed{\frac{1}{\dfrac{x_1-x_2}{\varepsilon(|x_1|+|x_2|)}-1}} \qquad (x_1 > x_2)$$

Si suponemos ahora que $\tilde{x}_2 > \tilde{x}_1$ y por consiguiente, que $x_2 > x_1$, tendríamos una nueva cota igual a la anterior, pero con $(x_2 - x_1)$:

$$\frac{1}{\dfrac{x_2-x_1}{\varepsilon(|x_1|+|x_2|)}-1}$$

Por tanto, la cota para la fórmula del error relativo será:

$$\frac{\left|\dfrac{1}{x_1-x_2}-\dfrac{1}{\tilde{x}_1-\tilde{x}_2}\right|}{\left|\dfrac{1}{x_1-x_2}\right|} \leq \frac{\varepsilon(|x_1|+|x_2|)}{|x_1-x_2|-\varepsilon(|x_1|+|x_2|)}$$

$$\boxed{\varepsilon < \frac{|x_1-x_2|}{|x_1|+|x_2|}}$$

COTA DEL ERROR RELATIVO.

* PROBLEMA 2.

Se considera una representación en coma flotante $x = m2^e$ con un exponente entero y una mantisa "m" $\in [1,2)$ $(1 \leq m < 2)$.

Indicar la mantisa y el exponente de:

a) $1/x$

b) $\sqrt{x}$

ⓐ $y = \dfrac{1}{x}$

$x = m2^e \quad , \quad e \in \mathbb{Z} \qquad 1 \leq m < 2$

$$\frac{1}{x} = \frac{1}{m} \cdot \frac{1}{2^e} = \underbrace{\frac{1}{m}}\cdot 2^{-e}$$

Podemos encontrar problemas en la mantisa porque debe $\in [1,2)$

El exponente no dará problema alguno.

¿ $m' = \dfrac{1}{m}$? ¿ Será $\dfrac{1}{m}$ la nueva mantisa ?

$\dfrac{1}{2} < \dfrac{1}{m} \leq 1 \longrightarrow$ Surgen problemas con los límites de la desigualdad.

CASO A.- $m = 1 \Longrightarrow \dfrac{1}{m} = 1$; $\boxed{m' = \dfrac{1}{m} \qquad e' = -e}$ $m' \in [1,2)$

$e' \in \mathbb{Z} \Longrightarrow \dfrac{1}{x} = \underset{\overset{\parallel}{1}}{m} \cdot 2^{-e} = 2^{-e}$

CASO B.- $m > 1 \Longrightarrow \dfrac{1}{2} < \dfrac{1}{m} \leq 1$ como $m \neq 1$ $\dfrac{1}{2} < \dfrac{1}{m} < 1$

Multiplico $x2 \Longrightarrow 1 < \dfrac{2}{m} < 2 \Longrightarrow 1 \leq m' < 2$

↳ Sirve como nueva mantisa

Ajustamos el exponente:

$$\frac{1}{x} = \frac{1}{m} \cdot 2^{-e} = \frac{2}{m} \cdot \frac{2^{-e}}{2} = \underbrace{\frac{2}{m}}_{m'} \cdot 2^{\overbrace{-e-1}^{e'}} \qquad \begin{array}{l} 1 \leq m' < 2 \\ e' = -e-1 \end{array}$$

Luego $\frac{1}{x} = m' . 2^{e'}$

$$\boxed{\begin{array}{c} m' = \frac{2}{m} \\ \\ e' = e - 1 \end{array}}$$

(b) $y = \sqrt{x}$

$x = m 2^e \quad ; \quad 1 \leq m < 2 \quad , \quad e \in \mathbb{Z}$

$$\sqrt{x} = x^{1/2} = m^{1/2} . (2^e)^{1/2} = m^{1/2} . 2^{e/2}$$

Podemos encontrar problemas en el exponente porque tiene que ser entero. La mantisa en este caso no da problemas.

CASO A.- e par

$$e = 2p \quad , \quad p \in \mathbb{Z} \implies \sqrt{x} = m^{1/2} . 2^{e/2} = m^{1/2} . 2^p \qquad p \in \mathbb{Z}$$

$$1 \leq m < 2 \implies \underset{\overset{\shortparallel}{1}}{\sqrt{1}} \leq m^{1/2} < \underset{1'4 < 2}{\sqrt{2}} \implies 1 \leq m^{1/2} < 2$$

$$\boxed{\begin{array}{c} m' = \sqrt{m} \\ e' = \frac{e}{2} \end{array}}$$

EL SIGNO DEL NÚMERO NOS DA IGUAL PORQUE VA EN UN BIT APARTE.

CASO B.- e impar

$e = 2p + 1 \quad , \quad p \in \mathbb{Z}$

$$\sqrt{x} = \sqrt{m} . 2^{e/2} = \sqrt{m} . 2^{p+\frac{1}{2}} = \sqrt{m} . 2^p . 2^{1/2} = \sqrt{2m} . 2^p$$

¿Me sirve como m'?

$$1 \leq m < 2 \xrightarrow{\times 2} 2 \leq 2m < 4 \xrightarrow{\sqrt{\ }} \underset{\substack{\shortparallel \\ 1'4 > 1}}{\sqrt{2}} \leq \sqrt{2m} < 2$$

$$1 \leq \sqrt{2} \leq \sqrt{2m} < 2 \implies 1 \leq \sqrt{2m} < 2$$

$$\boxed{\begin{array}{l} m' = \sqrt{2m} \\ e' = \dfrac{e-1}{2} \end{array}}$$

Ésta es la nueva representación.

* EJERCICIO (pendiente del día anterior).

Tenemos un número x que lo aproximo a $\tilde{x}$ para representarlo.

$x = \pm (1+m)2^e$ con m de más de 52 bits.

$\tilde{x} = \pm (1+\tilde{m})2^{\tilde{e}}$ con $\tilde{m}$ de 52 bits.

$e = \tilde{e}$ (Nos fijamos sólo en la mantisa)

1 bit	11 bits	52 bits
		$a_1 a_2 \quad\quad a_{52}$
Signo	Exponente	Mantisa

$a_i \in \{0,1\}$

$$1 \leq m = 1 + a_1\left(\frac{1}{2}\right) + a_2\left(\frac{1}{2}\right)^2 + \cdots + a_{52}\cdot\left(\frac{1}{2}\right)^{52} \leq 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^{52} =$$

$$= 1 + \frac{\frac{1}{2} - \overset{<1}{\frac{1}{2^{52}}}}{\frac{1}{2}} = 1 + \frac{1}{2^{52}} \leq 1 + \boxed{\ } < 2$$

$$x \rightarrow a_1 a_2 \ldots a_{52} \,\Big|\, \underbrace{a_{53} \ldots}_{\text{Error cometido en la mantisa al truncar}}$$

TRUNCAMOS

$$\tilde{x} \rightarrow a_1 a_2 \ldots a_{52}$$

Error cometido en la mantisa al truncar $\rightarrow a_{53}\left(\frac{1}{2}\right)^{53} + a_{54}\left(\frac{1}{2}\right)^{54} + \cdots + \leq \left(\frac{1}{2}\right)^{53} + \left(\frac{1}{2}\right)^{54} + \cdots =$

$$= \sum_{n \geq 53} (1/2)^n = \frac{(1/2)^{53}}{1 - \frac{1}{2}} = \frac{1}{2^{52}} = 2^{-52}$$

Serie geométrica de razón $1/2$

Sea $x = m 2^e \longrightarrow \tilde{x} = \tilde{m} 2^{\tilde{e}}$ con $\tilde{e} = e$

$$\left.\begin{aligned} x &= m\,2^e \\ \tilde{x} &= \tilde{m}\,2^e \end{aligned}\right\}$$ Restamos y obtenemos:

$x - \tilde{x} = (m - \tilde{m}) . 2^e$ Tomamos valores absolutos...

ERROR ABSOLUTO $\rightarrow |x - \tilde{x}| = |m - \tilde{m}| . 2^e$

ERROR RELATIVO $\rightarrow \dfrac{|x - \tilde{x}|}{|x|} = \dfrac{|m - \tilde{m}|\,\cancel{2^e}}{m\,\cancel{2^e}} = |m - \tilde{m}| \leq 2^{-52}$ (*)

$\hookrightarrow$ Tomamos $m = 1$

Siendo ÈPSILÓN DE LA MÁQUINA una cota superior de δ donde:

$$\tilde{x} = x(1 + \delta) = x + x\delta$$

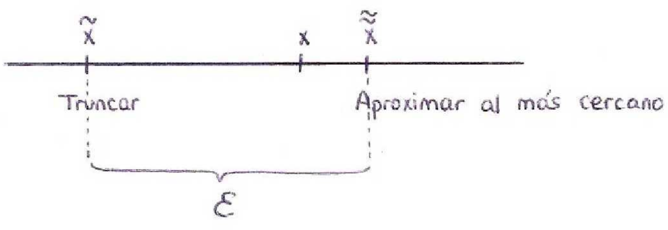
$$\delta = \frac{\tilde{x} - x}{x} \quad ; \quad |\delta| = \text{ERROR RELATIVO}$$

Luego, Èpsilón de la máquina:

$$\boxed{\varepsilon \leq 2^{-52}}$$ en el estándar IEEE.

$\hookrightarrow$ TRUNCAR

¿Y si en lugar de truncar, redondeo?



LOS ERRORES SE VAN ACUMULANDO. PARA DISMINUIR EL ERROR ES MEJOR SUMAR/RESTAR VARIAS VECES Y NO MULTIPLICAR/DIVIDIR.

$\frac{1}{2} . 2^{-52} = 2^{-53} \rightarrow$ Cota de ε $\rightarrow$ REDONDEAR

(Como son cantidades muy pequeñas, 2^{-52} y 2^{-53} son casi iguales).